Biochem. J. (1990) 272, 493–497 (Printed in Great Britain)

Characterization of the human lysosomal α-glucosidase gene

Lies H. HOEFSLOOT, Marianne HOOGEVEEN-WESTERVELD, Arnold J. J. REUSER* and Ben A. OOSTRA
MGC-Department of Cell Biology and Genetics, Erasmus University, Rotterdam, The Netherlands

The gene coding for human lysosomal α-glucosidase was cloned and its structure was determined. The gene is approx. 20 kb long, and contains 20 exons. The first exon is non-coding. The coding sequence of the putative catalytic site domain is interrupted in the middle by an intron of 101 bp. This intron is not conserved in the highly similar region of the human and rabbit isomaltase genes. The promoter region was defined by a CAT assay and the start of the mRNA was determined by primer extension. The promoter has features characteristic of a 'housekeeping' gene. The GC content is high (80%) and distinct TATA and CCAAT motifs are lacking. Two potential binding sites for the AP-2 transcription factor are present. Four potential Sp-1 binding sites are located downstream of the 5′ end of the mRNA.

INTRODUCTION

Lysosomal α-glucosidase (acid α-glucosidase; glucan 1,4-α-glucosidase; EC 3.2.1.3) is essential for the degradation of lysosomal deposits of glycogen. Inherited enzyme deficiency leads to lysosomal glycogen storage disease type II (glycogenosis type II; Pompe disease) (Hers, 1963). Several distinct abnormalities in enzyme synthesis and post-translational modification have been discovered in the various clinical phenotypes of this disease (Reuser *et al.*, 1985, 1987; Van der Ploeg *et al.*, 1988). The full-length cDNA coding for acid α-glucosidase has been cloned (Hoefsloot *et al.*, 1988) and expressed in mammalian cells (Hoefsloot *et al.*, 1990*a*). The cDNA-encoded enzyme was shown to have the same characteristics as the endogenous acid α-glucosidase of human fibroblasts with respect to intracellular transport, post-translational modification and function. One of the remarkable features of acid α-glucosidase is its sequence similarity with both subunits of the intestinal sucrase–isomaltase enzyme complex (Hoefsloot *et al.*, 1988). Based on this similarity, the catalytic site of acid α-glucosidase was assigned tentatively (Quaroni & Semenza, 1976; Hunziker *et al.*, 1986). In the present report we describe the organization of the acid α-glucosidase gene and the characteristic features of the promoter region. The gene structures around the putative catalytic sites of acid α-glucosidase and isomaltase are compared.

EXPERIMENTAL

Isolation of genomic clones

A human genomic EMBL-3 library (CML-0; De Klein *et al.*, 1986) was screened with a full-length human acid α-glucosidase cDNA, clone pSHAG2 (Hoefsloot *et al.*, 1990*a*). Hybridizing restriction fragments of the isolated phage clones were subcloned in appropriate sites of either pTZ18 or M13mp18/mp19 (Pharmacia, Uppsala, Sweden). The inserts were sequenced using the T7 polymerase sequencing kit according to the instructions of the manufacturer (Pharmacia). The M13 universal primer, or primers complementary to the cDNA, were used.

Southern blotting

DNA was isolated from 10 ml blood samples obtained from unrelated Caucasians, using the high-salt extraction procedure (Miller *et al.*, 1988). Restriction enzyme digests were performed on 10–15 μg of DNA in the appropriate buffers. DNA fragments were separated on 0.8% (w/v) agarose gels and subsequently blotted on to nitrocellulose filters. Filters were hybridized with acid α-glucosidase cDNA using standard protocols (Sambrook *et al.*, 1989).

Polymerase chain reaction

DNA isolated from human control fibroblasts and rabbit liver was used as a template in a reaction mixture containing 100 pmol of each primer, 2 units of Amplitaq (Cetus), 50 mM-Tris/HCl (pH 8.3), 3.0 mM-$MgCl_2$, 25 mM-KCl, 200 μg of BSA/μl, 10% (v/v) dimethyl sulphoxide, 5 mM-β-mercaptoethanol, 17 mM-$(NH_4)_2SO_4$ and 0.1 mM of each dNTP. DNA fragments were amplified in 25 cycles (2 min of denaturation at 94 °C, 1.5 min of annealing at 57 °C, and 3 min of extension at 72 °C) using a Cetus DNA amplifier (Cetus, Emeryville, CA, U.S.A.). One-third of each reaction was analysed on a 2% (w/v) Nusieve/agarose gel.

CAT assay

The TK promoter of vector pBLCAT3 (Luckow & Schütz, 1987) was removed by digestion with *Bam*HI and *Bgl*II. Fragments of the 5′ region of the acid α-glucosidase gene were cloned in this vector as follows. A 2 kb *Stu*I–*Pvu*II fragment and a 325 bp *Sac*I–*Pvu*II fragment were subcloned in the *Sma*I site and the *Sac*I–*Sma*I sites respectively of pSP72 (Pharmacia). Using the *Bam*HI and *Bgl*II site on either side, the insert was retrieved from this vector and cloned in the corresponding sites of pBLCAT3 in sense orientation. A series of 5′ deletion clones derived from the 2 kb fragment in pBLCAT3 was generated by exonuclease III digestion. COS-1 cells were transfected with the CAT constructs as described before (Hoefsloot *et al.*, 1990*a*). Cells were harvested 72 h after transfection and lysed by repeated freeze–thawing in 0.25 M-Tris/HCl (pH 7.8). A 10000 *g* supernatant was prepared and endogenous acetylases were inactivated by incubation for 10 min at 60 °C. CAT activity was determined according to Gorman *et al.* (1982).

Sequence analysis of promoter region

Several restriction fragments derived from the 5′ end of the acid α-glucosidase gene were subcloned in M13mp18/mp19 and

Abbreviation used: CAT, chloramphenicol acetyltransferase.

* To whom correspondence and reprint requests should be addressed, at: Department of Cell Biology and Genetics, Erasmus University, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands.

The nucleotide sequence data in this paper have been submitted to the EMBL/GenBank/DDBJ Nucleotide Sequence Databases under accession nos. X55079–X55098.

sequenced in both directions. In addition, relevant exonuclease III-generated CAT constructs were sequenced from their 5′ end using double-stranded plasmid DNA and the M13 universal primer.

Primer extension

RNA was isolated from human fibroblasts using the method of Schreiber *et al.* (1989). Synthetic RNA was made as described previously (Melton *et al.*, 1984). Oligonucleotides were end-labelled using [γ-^{32}P]dATP and polynucleotide kinase. Radiolabelled oligonucleotide (10^6 c.p.m.) was hybridized for 8–12 h at 32 °C to 100 μg of RNA in a 30 μl reaction mixture containing 40 mM-Pipes (pH 6.4), 0.4 M-NaCl, 1 mM-EDTA (pH 8.0) and 80 % (v/v) formamide. The extension reaction was carried out according to Sambrook *et al.* (1989) and products were analysed on a 10 % polyacrylamide gel with 1 % cross-linking.

RESULTS

Gene structure

Eight overlapping λ clones hybridizing with acid α-glucosidase cDNA were isolated from a human genomic library. Together, these clones span a region of more than 33 kb (Fig. 1). All hybridizing sequences were contained within three contiguous *Bgl*II fragments of 10.5, 8.5 and 14 kb, which were subcloned in the *Bam*HI site of pTZ18. A partial restriction map was constructed and fragments containing exon sequences were identified using oligonucleotides corresponding to various cDNA regions. All exons and flanking regions were sequenced completely. The intron–exon boundaries were established by comparing the cDNA and genomic sequences. Using this strategy the spatial distribution of the exons and introns of the acid α-glucosidase gene was obtained (Fig. 1).

The gene contains 20 exons. The start codon of acid α-glucosidase is localized near the 5′ end of exon 2. Therefore exon 1 is non-coding. The stop codon is situated near the 5′ end of exon 20. All intron–exon boundaries conform to the 'GT/AG' rule, except for the splice donor site of exon 19, which includes a GC instead of a GT (Table 1). All three codon phases were encountered at the intron–exon boundaries.

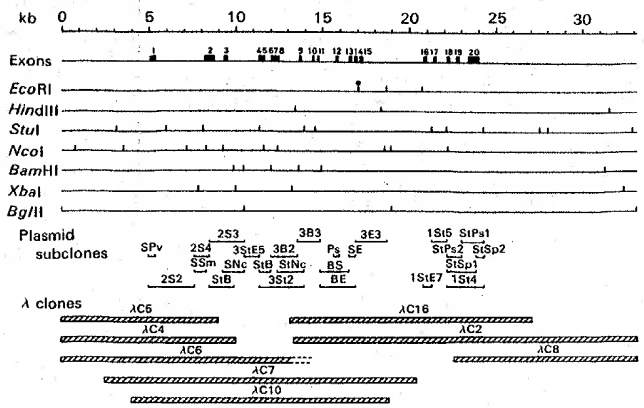

Fig. 1. Organization of the gene coding for acid α-glucosidase

A partial restriction map is given. The isolated phage clones and the plasmid subclones used for sequence analysis are indicated. The polymorphic *Eco*RI site is marked with an asterisk. The exons are represented by black boxes.

Table 1. Nucleotide sequences of the exon–intron boundaries

Exon and intron sites are given in numbers of base pairs. Introns 4, 6, 7, 10 and 13 were sequenced, and the exact size is indicated. The sizes of the other introns are based on the restriction map. Exon sequences are in upper case letters, intron sequences are in lower case. cDNA position refers to the numbering of the cDNA sequence as deposited in the EMBL/Genbank/DDBJ Nucleotide Sequence Databases under accession number Y00839. Codon phase 0 interrupts the coding sequence between two codons, phase I after the first nucleotide and phase II after the second nucleotide of a triplet.

Exon	Exon size	cDNA position of exon	5′ splice donor	3′ splice acceptor	Intron size	Codon phase
1	>187	→187	CGGgtagag	tctcccgcagGCC	2800	
2	578	188–765	ACGgtgggc	tctcttctagATC	600	0
3	146	766–911	GCTgtgagt	tgtcccgcagGCT	1850	II
4	166	912–1077	ACGgtacag	gcatgtccagCCC	84	0
5	97	1078–1174	TGGgtaagc	tccctccagATG	350	I
6	120	1175–1294	TGGgtaggg	tggcctgcagGAT	80	I
7	119	1295–1413	CTGgtgagt	tgtgctgcagGAC	88	0
8	132	1414–1545	GTGgtgtgt	ctcttcccagGAT	1120	0
9	111	1546–1656	AAGgtaggg	cgttgtccagGTA	670	0
10	114	1657–1770	ATTgtaagt	tctcttgcagGAC	101	0
11	85	1771–1855	CTGgtcagc	cctcttccagGGG	820	I
12	118	1856–1973	CAGgtgagg	accaccccagGGC	600	II
13	134	1974–2107	CAGgtaagc	gccctcccagAAA	139	I
14	152	2108–2259	CTGgtaggg	tgccctgcagCCC	190	0
15	149	2260–2408	GGAgtgagt	cccctttgcagGTT	3600	II
16	142	2409–2550	ACGgtgagt	ctccctccagGTG	450	0
17	150	2551–2700	CAGgtacct	ccctttccagGGC	650	0
18	165	2701–2865	AATgtgagt	ctcggcccagAAC	350	0
19	153	2866–3018	AAGgcaaga	ctctttccagGTC	550	0
20	606	3019–3624				
Consensus:			$\begin{matrix}A\\C\end{matrix}AGgt\begin{matrix}a\\g\end{matrix}aga$	$\left(\begin{matrix}t\\c\end{matrix}\right)_n\begin{matrix}c\\t\end{matrix}agG$		

Catalytic site domain

Based on the sequence similarity between acid α-glucosidase and isomaltase, the aspartic acid residue encoded by nucleotides 1771–1773 was predicted to be the essential residue in the catalytic site of acid α-glucosidase (Hoefsloot *et al.*, 1988). Table 1 shows that an intron of 101 bp is localized between position 1770 and 1771 of the cDNA sequence, thus interrupting the coding sequence of the putative active site domain. To investigate the conservation of this intron during evolution, the corresponding domain of human and rabbit isomaltase was analysed using the polymerase chain reaction. One set of primers specific for acid α-glucosidase was chosen in exons 10 and 11. Using these primers for amplification of cDNA, the expected fragment of 190 bp was obtained (Fig. 2). Amplification of genomic DNA with the same primers resulted in the expected longer fragment. A second set of primers was chosen to analyse the corresponding domain in human and rabbit isomaltase using the published cDNA sequences (Hunziker *et al.*, 1986; Green *et al.*, 1987). The amplified cDNA and genomic fragments of isomaltase were exactly the same size. Thus the sequence coding for the catalytic site of human and rabbit isomaltase is not interrupted by an intron.

Promoter region and transcription initiation site

To define the promoter region, two genomic fragments of different lengths were subcloned in front of the bacterial CAT gene (Fig. 3). The longer 2 kb fragment (*Stu*I–*Pvu*II) did promote CAT activity in transfected COS cells. No activity was detected with a construct containing the smaller 325 bp fragment (*Sac*I–*Pvu*II). To determine the position of the promoter region

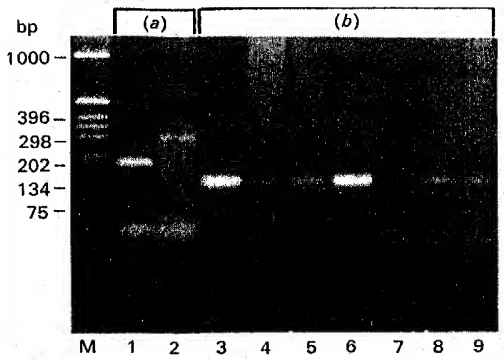

Fig. 2. Polymerase chain reaction analysis of acid α-glucosidase and isomaltase gene structure around the catalytic site

(*a*) Polymerase chain reaction with acid α-glucosidase-specific primers. These were 5′-TATGGCCCGGGTCCACTGCC (sense) and 5′-CAGGCACGTAGGGTGGGTTCTC (anti-sense). (*b*) Polymerase chain reaction with isomaltase (human and rabbit)-specific primers. These had the sequences 5′-TGATTTCACTA-ATCCAAACTGCA (sense) and 5′-CATTACATCCTTTTGTTG-AACCT (anti-sense). Templates were human acid α-glucosidase cDNA (lanes 1 and 7), human genomic DNAs (lanes 2, 4 and 5), human isomaltase cDNA (lanes 3 and 6) (Green *et al.*, 1987), rabbit sucrase–isomaltase cDNA (lane 8) (Hunziker *et al.*, 1986) and rabbit liver DNA (lane 9). Lane M contains markers. Fragment lengths are given in base pairs.

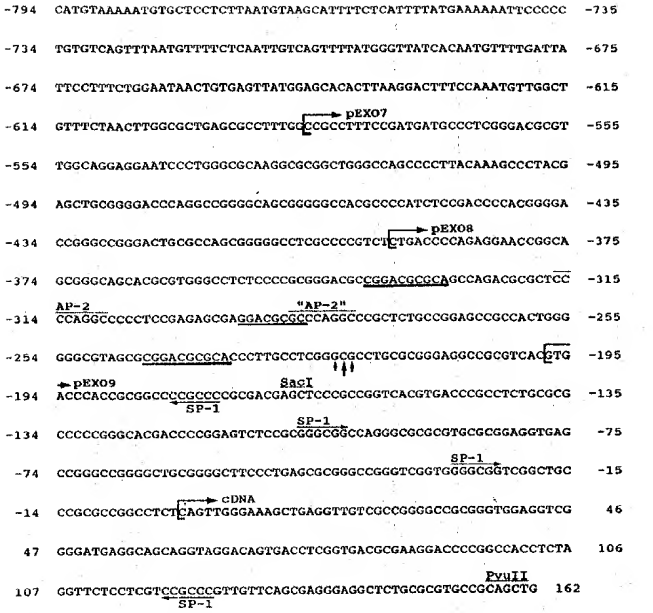

Fig. 4. Sequence of the 5′ end of the acid α-glucosidase gene

The 5′ end of the exonuclease III-generated clones and the beginning of the longest cloned cDNA are indicated, as well as some restriction sites. Thick line, 10(8) bp repeat; AP-2, putative binding sites for the AP-2 transcription factor; SP-1, putative binding sites for the SP-1 transcription factor. Arrows indicate the transcription initiation site.

more precisely, the 2 kb fragment was shortened from the 5′ end by using exonuclease III (Fig. 3). Transient expression of these constructs in COS cells showed that only the shortest construct (pEXO9) has lost promoter activity. The other constructs were equally effective in expressing CAT activity. Thus the promoter region does not extend upstream of clone pEXO8. The nucleotide sequence comprising the 5′ end of the acid α-glucosidase gene is given in Fig. 4. The start points of the exonuclease-generated clones, as well as the beginning of the longest cloned cDNA, are indicated.

The 5′ end of the acid α-glucosidase mRNA was determined by primer extension of a 28-mer oligonucleotide complementary to positions −71 to −98 (Fig. 4). Using this oligonucleotide the longest fragment obtained had a length of 150–152 nucleotides (Fig. 5). This places the transcription initiation site of acid α-glucosidase between positions −220 and −222. In addition, a smaller fragment of 134 nucleotides was detected, which could be explained by premature termination caused by secondary structures. Smaller fragments than expected were also obtained using RNA from this region synthesized *in vitro* (results not shown).

The promoter region defined by the CAT assay and primer

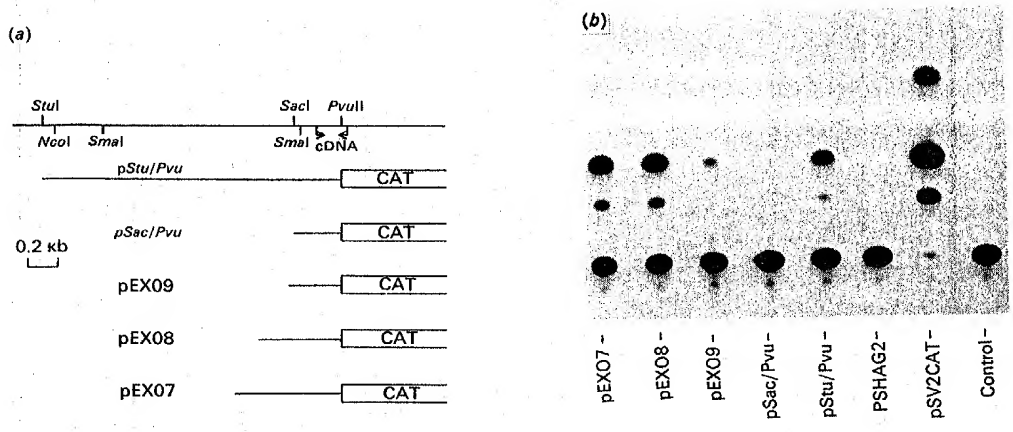

Fig. 3. CAT constructs and their expression in COS-1 cells

(*a*) Restriction map of the 5′ end of the acid α-glucosidase gene. The various constructs used in the CAT assay are indicated. The arrows in the top line indicate the beginning of the longest cloned cDNA and the end of exon 1. (*b*) CAT assays with lysates from COS-1 cells transfected with CAT constructs. Mock-transfected cells serve as a control. pSV2CAT, SV40 promoter in front of the CAT gene.

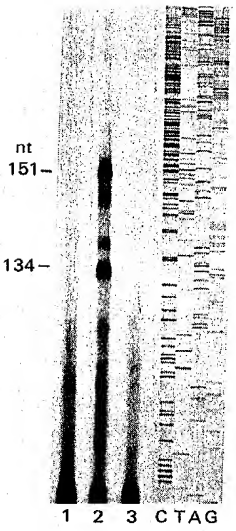

Fig. 5. Primer extension with an oligonucleotide complementary to positions −71 to −98

A sequence reaction from a fragment of the 5′ region was used as a size marker. Lane 1, primer only; lane 2, primer extension with 100 μg of total RNA isolated from human fibroblasts; lane 3, reaction with 100 μg of tRNA. Numbers on the left indicate lengths of fragments in nucleotides.

extension does not contain a typical CCAAT box or a TATA-resembling motif (Fig. 4). A potential AP-2 binding site with a perfect match to the consensus sequence (Mitchell & Tjian, 1989) is located at positions −316 to −309, and a second site with one mismatch is located at positions −287 to −280. There are several direct repeats, the longest of which is found at positions −338 to −329 and −244 to −235. The middle eight base pairs of this repeat recur at positions −293 to −286. The sequence (Fig. 4) includes four potential Sp-1 binding sites (Dynan, 1986; Mitchell & Tjian, 1989), two in sense and two in anti-sense orientation. However, these are all located in the untranslated region of the acid α-glucosidase mRNA. The G+C content is 80% and the observed/expected ratio of the CpG dinucleotide is 0.9. The combined features are typical of those for the promoter of a housekeeping gene (Dynan, 1986).

Table 2. DNA polymorphisms

cDNA position refers to the numbering of the cDNA sequence as deposited in the EMBL/GenBank/DDBJ Nucleotide Sequence Databases under accession number Y00839.

cDNA position	cDNA	Genomic	Amino acid alteration
543	C	T	–
815	G	A	Arg → His (conservative)
887	A	G	His → Arg (conservative)
1423	C	T	Arg → Trp (hydrophobic)
1800	A	G	–
2772	A	G	–
3217/3218	–	G	Non-coding
3305	G	C	Non-coding
3496	T	G	Non-coding

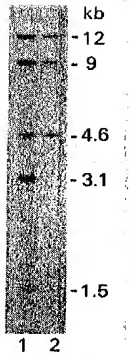

Fig. 6. *Eco*RI polymorphism in the acid α-glucosidase gene

*Hin*dIII–*Eco*RI-double-digested DNA was analysed by Southern blotting using the full-length acid α-glucosidase cDNA as a probe. The length of hybridizing fragments is indicated in kb.

DNA polymorphisms

Sequence comparison of all exonic DNA sequences and the previously published acid α-glucosidase cDNA revealed several differences. Some of these differences appeared to be artificial and were found to be caused by misinterpretation of the cDNA sequence data. Others were identified as single base pair polymorphisms. These are listed in Table 2. The corrected cDNA sequence has been submitted to the EMBL/GenBank/DDBJ Nucleotide Sequence Databases (accession numbers X55079–X55098).

According to the restriction map in Fig. 1, three genomic *Eco*RI fragments are expected to hybridize with acid α-glucosidase cDNA. The 1.5 kb fragment, however, containing exon 15, was not detected in previous Southern blot hybridizations (Hoefsloot *et al.*, 1988). To investigate whether the 5′ *Eco*RI site of the 1.5 kb fragment (marked with an asterisk) is polymorphic, DNA of 11 unrelated individuals was analysed. To facilitate the interpretation of the results, the DNA was double-digested with *Hin*dIII and *Eco*RI. In case the *Eco*RI site is present, the 4.6 kb *Hin*dIII fragment (Fig. 1) was cut into two smaller fragments of 3.1 and 1.5 kb. An example is given in Fig. 6. Heterozygosity for the *Eco*RI polymorphism (Fig. 6, lane 1) was detected in three out of eleven cases. Sequence analysis of both alleles showed that the polymorphism is based on the variable presence of a thymidine residue in the GAA(T)TC *Eco*RI recognition sequence.

DISCUSSION

The gene coding for human acid α-glucosidase contains 20 exons and 19 introns spread over a distance of 20 kb. The sizes of the exons and introns are not unusual for eukaryotic genes (Hawkins, 1988). The first intron is located within the 5′ untranslated region, and the first exon is therefore non-coding. The ATG start codon is located 33 bp from the beginning of the second exon. The untranslated part of the first coding exon of vertebrate genes is generally short, and rarely exceeds 40 nucleotides (Hawkins, 1988). It has been suggested that introns demarcate structural and/or functional domains of proteins (Gilbert, 1985). For instance, a correlation between structural domains and intron–exon organization was postulated for lysosomal acid phosphatase (Geier *et al.*, 1989). Furthermore, the proteolytic cleavage site used in the maturation of the α-chain of lysosomal hexosaminidase is located at the beginning of an exon (Proia, 1988). Little information is available yet on the structural domains of acid α-glucosidase. However, the signal

peptide, the pro-sequence of the acid α-glucosidase precursor and first 61 amino acids of the 76 kDa mature enzyme are all encoded in the same exon (exon 2). It is also notable that the coding information for the putative catalytic site domain of acid α-glucosidase is interrupted by an intron. Considering the strong sequence similarity between acid α-glucosidase and isomaltase, it is surprising that no intron is present at the same site in the human and rabbit isomaltase gene.

All splice junctions conform to the 'GT/AG' rule, except for the splice donor site of exon 19, which has GC instead of GT. Such a splice donor site is very rare, but has been described for human and rodent adenine phosphoribosyltransferase genes (Broderick *et al.*, 1987), duck (Erbil & Niessing, 1983) and chicken (Dodgson & Engel, 1983) α-globin genes, and the mouse αA-crystallin (King & Piatigorsky, 1983) gene.

The transcription initiation site was determined by primer extension, and was found to be located approx. 220 bp upstream from the longest cloned cDNA. This is 440 bp in front of the ATG start codon. The transcription initiation site is properly positioned within the limits of the promoter region, as determined by the various constructs used in the CAT assay. The *Sac*I–*Pvu*II fragment located 3′ of the transcription initiation site lacks promoter activity. A genomic fragment starting 175 bp upstream of the transcription initiation site (clone pEXO8) has full promoter activity. The characteristics of this region are typical for the promoter of a housekeeping gene. The G+C content is high (80 %) and the CpG dinucleotide is not depleted, meeting the requirements for a CpG island (Gardiner-Garden & Frommer, 1987). Sequences resembling TATA motifs are absent. The CCACT sequence at positions −262 to −258 is located too close to the proposed transcription initiation site to function as a CCAAT box (Breathnach & Chambon, 1981). The promoter regions of a few other lysosomal enzyme genes have been studied (Proia & Soravia, 1987; Bishop *et al.*, 1988; Neote *et al.*, 1988; Geier *et al.*, 1989) and all except one (glucocerebrosidase; see Horowitz *et al.*, 1989) seem to have a promoter characteristic of a housekeeping gene. The presence of one, possibly two, putative AP-2 binding sites (Mitchell & Tjian, 1989) in the promoter region of acid α-glucosidase is remarkable, since the AP-2 transcription factor confers inducibility of gene expression by cyclicAMP and phorbol esters (Imagawa *et al.*, 1987). Whether the AP-2 binding sites are relevant for acid α-glucosidase expression remains to be determined. In the 5′ flanking sequences of the hexosaminidase β-gene (Neote *et al.*, 1988) and the α-galactosidase gene (Bishop *et al.*, 1988), two and one putative AP-1 binding sites were found respectively. The AP-1 transcription factor confers inducibility by phorbol ester.

Several polymorphisms were found. Most were silent or conservative (Table 2). The only non-conservative difference concerns a C to T transition at nucleotide position 1423, leading to a substitution of arginine by tryptophan. Tryptophan-containing acid α-glucosidase was found to be transported to the lysosomes and to be catalytically active. The arginine-containing enzyme, however, did not mature, and was detected only in the endoplasmic reticulum and the Golgi complex, in a catalytically inactive form (Hoefsloot *et al.*, 1990*a*). The polymorphic *Eco*RI site is situated in intron 14. The recently reported *Xba*I polymorphism (Hoefsloot *et al.*, 1990*b*) is due to the variable presence of an *Xba*I site in the *Xba*I fragment containing exons 2 and 3 (Fig. 1). Both restriction fragment length polymorphisms can be used for diagnostic purposes.

Received 19 June 1990; accepted 20 July 1990

We thank Dr. Hans Galjaard for his continuous support, Dr. Gerard Grosveld for supplying the CML-0 library, Dr. Giorgio Semenza and Dr. Ned Mantei for the rabbit sucrase–isomaltase cDNA, and Dr. Dallas Swallow for the human isomaltase cDNA clone. We acknowledge Jeannette Lokker for typing the manuscript, and Pim Visser and Tom de Vries Lentsch for Figures and photography. Financial support was obtained from the Netherlands Organisation for Scientific Research (N.W.O.) and the Prinses Beatrix Fonds.

REFERENCES

Bishop, D. F., Kornreich, R. & Desnick, R. J. (1988) Proc. Natl. Acad. Sci. U.S.A. **85**, 3903–3907
Breathnach, R. & Chambon, P. (1981) Annu. Rev. Biochem. **50**, 349–383
Broderick, T. P., Schaff, D. A., Bertino, A. M., Dusch, M. K., Tischfield, J. A. & Stambrook, P. J. (1987) Proc. Natl. Acad. Sci. U.S.A. **84**, 3349–3353
De Klein, A., Hagemeijer, A., Bartram, C. R., Houwen, R., Hoefsloot, L., Carbonell, F., Chan, L., Barnett, M., Greaves, M., Kleihauer, E., Heisterkamp, N., Groffen, J. & Grosveld, G. (1986) Blood **68**, 1369–1375
Dodgson, J. B. & Engel, J. D. (1983) J. Biol. Chem. **258**, 4623–4629
Dynan, W. S. (1986) Trends Genet. **2**, 196–197
Erbil, C. & Niessing, J. (1983) EMBO J. **2**, 1339–1343
Gardiner-Garden, M. & Frommer, M. (1987) J. Mol. Biol. **196**, 261–282
Geier. C., Von Figura, K. & Pohlmann, R. (1989) Eur. J. Biochem. **183**, 611–616
Gilbert, W. (1985) Science **228**, 823–824
Gorman, C. M., Moffat, L. F. & Howard, B. H. (1982) Mol. Cell. Biol. **2**, 1044–1051
Green, F., Edwards, Y., Hauri, H.-P., Povey, S., Ho, M. W., Pinto, M. & Swallow, D. (1987) Gene **57**, 101–110
Hawkins, J. D. (1988) Nucleic Acids Res. **16**, 9893–9908
Hers, H. G. (1963) Biochem. J. **86**, 11–21
Hoefsloot, L. H., Hoogeveen-Westerveld, M., Kroos, M. A., van Beeumen, J., Reuser, A. J. J. & Oostra, B. A. (1988) EMBO J. **7**, 1697–1704
Hoefsloot, L. H., Willemsen, R., Kroos, M. A., Hoogeveen-Westerveld, M., Hermans, M. M. P., Van der Ploeg, A. T., Oostra, B. A. & Reuser, A. J. J. (1990*a*) Biochem. J. **272**, 485–492
Hoefsloot, L. H., Hoogeveen-Westerveld, M., Oostra, B. A. & Reuser, A. J. J. (1990*b*) Nucleic Acids Res., in the press
Horowitz, M., Wilder, S., Horowitz, Z., Reiner, O., Gelbart, T. & Beutler, E. (1989) Genomics **4**, 87–96
Hunziker, W., Spiess, M., Semenza, G. & Lodish, H. F. (1986) Cell **46**, 227–234
Imagawa, M., Chiu, R. & Karin, M. (1987) Cell **51**, 251–260
King, C. R. & Piatigorsky, J. (1983) Cell **32**, 707–712
Luckow, B. & Schütz, G. (1987) Nucleic Acids Res. **15**, 5490
Melton, D. A., Krieg, P. A., Rebagliati, M. R., Maniatis, T., Zinn, K. & Green, M. R. (1984) Nucleic Acids Res. **12**, 7035–7056
Miller, S. A., Dykes, D. D. & Polesky, H. F. (1988) Nucleic Acids Res. **16**, 1215
Mitchell, P. J. & Tjian, R. (1989) Science **245**, 371–378
Neote, K., Bapat, B., Dumbrille-Ross, A., Troxel, C., Schuster, S. M., Mahuran, D. J. & Gravel, R. A. (1988) Genomics **3**, 279–286
Proia, R. L. (1988) Proc. Natl. Acad. Sci. U.S.A. **85**, 1883–1887
Proia, R. L. & Soravia, E. (1987) J. Biol. Chem. **262**, 5677–5681
Quaroni, A. & Semenza, G. (1976) J. Biol. Chem. **251**, 3250–3253
Reuser, A. J. J., Kroos, M., Oude Elferink, R. P. J. & Tager, J. M. (1985) J. Biol. Chem. **260**, 8336–8341
Reuser, A. J. J., Kroos, M., Willemsen, R., Swallow, D., Tager, J. M. & Galjaard, H. (1987) J. Clin. Invest. **79**, 1689–1699
Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) Molecular Cloning, A Laboratory Manual, 2nd edn., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York
Schreiber, E., Matthias, P., Müller, M. M. & Schaffner, W. (1989) Nucleic Acids Res. **17**, 6419
Van der Ploeg, A. T., Bolhuis, P. A., Wolterman, R. A., Visser, J. W., Loonen, M. C. B., Busch, H. F. M. & Reuser, A. J. J. (1988) J. Neurol. **235**, 392–396